⑲ 日本国特許庁（JP）　　⑪ 特許出願公開

⑫ 公開特許公報（A）　　昭58—144961

⑤1Int. Cl.³	識別記号	庁内整理番号	⑬公開　昭和58年(1983) 8 月29日
G 06 F 13/04		7361—5B	
G 11 C 9/06		8219—5B	

発明の数　1
審査請求　有

（全 7 頁）

⑤4記憶領域管理方式

⑳特　　願　昭57—28643
㉒出　　願　昭57(1982) 2 月24日
⑫発 明 者　金子悟
川崎市中原区上小田中1015番地
富士通株式会社内
⑪出 願 人　富士通株式会社
川崎市中原区上小田中1015番地
⑭代 理 人　弁理士　京谷四郎

明　　細　　書

1. 発明の名称

記憶領域管理方式

2. 特許請求の範囲

2以上の階層を有する記憶装置の上位階層の記憶領域を管理する記憶領域管理方式において、アクセス対象を複数のグループに分け、グループ毎に占有できる記憶量の上限値を定め、アクセスがあったとき当該アクセス対象のデータが上記上位階層に登録されておらず、かつ上記上位階層の記憶装置に空の記憶単位がないとき、上記アクセス対象の属するグループの占有領域量が当該グループの上限値をこえている場合には当該グループ内で、当該グループの占有領域量が当該グループの上限値をこえていない場合には全体でLRU法により置き換えを行なって新たな記憶単位を得るようにしたことを特徴とする記憶領域管理方式。

3. 発明の詳細な説明

(1) 発明の技術分野

本発明は、階層的記憶装置に係り、特に上位階層の記憶装置を有効に使用することのできる記憶領域管理方式に関する。

(2) 従来技術と問題点

従来の記憶領域管理方式としては、LRU(Least Recently Used)方式や、FIFO(First In First Out)方式等が知られている。前者は最後にアクセスされた時刻が最も古いものから順に新たな割当てのために追い出していく方式であり、後者は登録された時刻の古いものから順に新たな割当てのために追い出していく方式である。これらは二つとも、全体を一つの管理グループで管理すると、性質の異なるものが一緒に管理されることになり、互いに悪影響を与える可能性がある。またFIFO方式ではアクセス頻度に関係なく、先に登録されたものから順に追い出されるためアクセスの仕方に動的に対応できない。

第1図はディスク・キャッシュ・サブシステム

を説明するブロック図、第2図はディスク・キャッシュ・サブシステムの処理を説明するフロー・チャート、第3図は従来のＬＲＵテーブルを示す図である。1はホスト・コンピュータ、2はディスク／ディスク・キャッシュ・制御装置、3はバッファ記憶装置、4はディスク制御アダプタ、5はディスク記憶装置を示す。ディスク・キャッシュ・サブシステムでは、下位記憶階層にディスク記憶装置5があり、上位記憶階層にバッファ記憶装置（通常ＭＯＳ、ＣＣＤ等の半導体記憶装置）3がある。バッファ記憶装置3にはディスク記憶装置5のデータのコピーを格納する。記憶領域の管理単位は「トラック」であるが、固定長形式ディスク記憶装置では「ブロック」の場合もある。このようなディスク・キャッシュ・サブシステムを第2図によって説明する。ホスト・コンピュータ1があるトラックにアクセスしようとしたときは、次のように処理される。

① ディスク／ディスク・キャッシュ制御装置2において、エクステント・テーブル（バッファ記憶装置3に格納する対象となるトラック・アドレスを記入したテーブル）にあるものか否かを判定する。

Ｎｏの場合は従来通りディスク制御アダプタ4を介してディスク記憶装置5からデータを転送し、Ｙｅｓの場合は②の処理を行う。

② アクセス・アドレスがバッファ・テーブル（トラック・アドレスとバッファ記憶アドレスとの対応表）に登録されているかを判定する。

Ｙｅｓの場合は③、Ｎｏの場合は④の処理を行う。

③ バッファ記憶装置3からデータを転送する。

④ ディスク制御アダプタ4を介してディスク記憶装置5からバッファ記憶装置3へステージング（下位階層の記憶装置から上位階層の記憶装置へデータを複写）する。

ステージングするときは、バッファ記憶装置3にステージングするトラックのために新しい領域を割当てなければならないが、バッファ記憶装置3に空領域がなければ、どれかのトラックを追い出してその領域を新しいトラックのために提供する。そのトラックを追い出すときに使われるのがＬＲＵである。ＬＲＵテーブルには、従来、第3図のようにドライブ／トラック・アドレスが収納され、古いものは下位にあって、上へ順に新しくなっている。

ディスク・キャッシュ・サブシステムでは新規のトラックをＬＲＵテーブルに登録するのはステージングの時であり、この時ＬＲＵテーブルの最上位に登録される。今までの最上位のトラックは一つ下がる。途中このトラックにアクセスがあると、その時点でこのトラックは再び最上位に移される。アクセスが無ければ、新規トラックの登録の度に順位が下がっていく。新規トラックのための領域を割当てる時には最下位にあるトラックが機械的に選ばれ登録が抹消される。この操作と新規登録の操作とで途中のトラックの順位は一斉に一つずつ下がる。従ってある期間アクセスされないトラックはいずれ最下位となり、次の割当て時に追い出されることになる。つまりＬＲＵ方式は、アクセスの時間的局所性を仮定して、将来のアクセスを予測するものである。（長い間アクセスされないものは、将来もアクセスされない、逆に最近アクセスがあったものは近い将来アクセスされる可能性が高い）

従来のＬＲＵでは、この構造が全体で一つとなっている。その場合の問題点としては、アクセス間隔と将来のアクセスの可能性の関係は、データセットの性質、システムの環境、等によって均一ではないということである。

従って全部のトラックを均等に扱ってしまうと、将来アクセスされる可能性のあるトラックが追い出されたり、すぐに追い出してもいいトラックが長く占有したりするという現象が起きる。これに対処する方法としては、ＬＲＵテーブルへの登録位置を優先順位によって変えたり、いくつかのトラックを集めたグループを作り、グループ毎にＬＲＵテーブルを定義するというような方法があるが、前者は「優先順位」という表現のとおり、優先度の低いトラックは高いトラックの影響を被っ

てしまうし、優先順位のつけ方も難しい。後者はグループ毎に領域量をキチンと割振ってしまうと、他グループの領域に余りがあってもＬＲＵが働くことになり無駄が生ずるおそれがある。

(3) 発明の目的

本発明は、上記の問題点を解決するものであって、アクセス対象をアクセスの性質によってグループ分けし、ある条件によってグループ毎にＬＲＵを働かせることにより、他グループへの影響を小さくすることを目的とするものである。

(4) 発明の構成

上記目的を達成するため、本発明の記憶領域管理方式は、２以上の階層を有する記憶装置の上位階層の記憶領域を管理する記憶領域管理方式において、アクセス対象を複数のグループに分け、グループ毎に占有できる記憶量の上限値を定め、アクセスがあったとき当該アクセス対象のデータが上記上位階層に登録されておらず、かつ上記上位階層の記憶装置に空の記憶単位がないとき、上記アクセス対象の属するグループの占有領域量が当該グループの上限値をこえている場合には当該グループ内で、当該グループの占有領域量が当該グループの上限値をこえていない場合には全体でＬＲＵ法により置き換えを行って新たな記憶単位を得るようにしたことを特徴とするものである。

(5) 発明の実施例

以下、本発明を図面を参照しつつ説明する。

第４図は本発明に使用されるハードウェアの一実施例を示すブロック図、第５図は本発明の一実施例を説明するフローチャートである。第４図において、６，11ないし18と22はアンド・ゲート、７はレジスタ、８はカウンタ、９と19は比較回路、10と21はオア・ゲート、20はアドレス・レジスタ、23はＬＲＵテーブルを示す。アンド・ゲート６の入力端子の一方にはクロック信号ＣＬＫが、他方には上限値セット信号が供給され、出力端子はレジスタ７のセット端子に接続される。レジスタ７の入力端子には割当上限値が入力され、出力端子は比較回路９の一方の入力端子に接続される。カウンタ８のカウント・アップ端子には割当て信号が供給され、カウント・ダウン端子には解放信号が供給され、出力端子は比較回路９の他方の入力端子に接続される。比較回路９の出力端子はオア・ゲート10の入力端子およびグループ"０"のアンド・ゲート11の一方の入力端子に接続される。以上のアンド・ゲート６、レジスタ７、カウンタ８と比較回路９を有する回路構成はグループ"０"に係るものであるが、グループ1ないし３についても同様の回路構成であり、それぞれの比較回路の出力端子がオア・ゲート10の入力端子および対応するアンド・ゲート12ないし14の一方の入力端子に接続されている。アンド・ゲート11の他方の入力端子にはグループＩＤ〝００〟が入力され、アンド・ゲート12の他方の入力端子にはグループＩＤ〝０１〟が入力され、アンド・ゲート13の他方の入力端子にはグループＩＤ〝１０〟が入力され、アンド・ゲート14の他方の入力端子にはグループＩＤ〝１１〟が入力され、アンド・ゲート11ないし14の出力端子は比較回路19の入力端子に接続される。オア・ゲート10の出力端子はアンド・ゲート15と18の一方の入力端子に接続される。アンド・ゲート15の他方の入力端子にはＬＲＵ起動信号が入力され、出力端子はアンド・ゲート16の他方の入力端子に接続される。アンド・ゲート18の他方の入力端子にはタイミング信号Ｔが供給され、出力端子は比較回路19のセット端子に接続される。比較回路19の入力端子にはＬＲＵテーブル23からひかれたグループＩＤが入力され、出力端子はオア・ゲート21の入力端子に接続される。オア・ゲート21の出力端子はアンド・ゲート22の他方の入力端子に接続される。アンド・ゲート16の一方の入力端子にはタイミング信号Ｔが入力され、出力端子はアドレス・レジスタ20のアドレス・セット端子に接続される。アンド・ゲート17の一方の入力端子にはＬＲＵテーブル23の中のアドレス信号ｑが入力され、他方の入力端子にはＬＲＵ起動信号が入力され、出力端子はアドレス・レジスタ20の入力端子に接続される。アドレス・レジスタ20の出力端子はアンド・ゲート22の一方の入力端子およびＬＲＵテーブル23の読み出し部に接続され

る。ＬＲＵテーブルは第５図に示すようにドライブ／トラック・アドレスとグループＩＤが収納され、割当上限値は第６図に示すようにグループＩＤ毎に設定される。

以上の構成により本発明の一実施例を第７図のフロー・チャートに沿って説明する。占有できる記憶領域量（割当上限値）Ｌ(i)はレジスタ７、占有領域量Ａ(i)は割当、解放に従い増、減するカウンタ８（ｉはグループｉの意味）から読み出され、両者は比較回路９において比較される。そこで、ホスト・コンピュータ１からアクセスがあったが、第２図のフロー・チャートの②の処理においてバッファ・テーブルに登録されておらず、空領域がないときは次のように処理を行う。

① ＬＲＵによる置き換え処理が開始される。

② 各グループのカウンタ８の出力Ａ(i)、即ち占有領域量が割当上限値Ｌ(i)に達しているかどうかを比較回路９により判定する。Ｙｅｓの場合は③を経て⑤、Ｎｏの場合は④を経て⑤の処理をする。

③ オア・ゲート10の出力とＬＲＵ起動信号によりアンド・ゲート15の条件が成立し、アンド・ゲート15の出力とタイミング信号Ｔによりアンド・ゲート16の条件が成立するのでアドレス・レジスタ20にアドレス信号ｑがセットされる。アドレス・レジスタ20によりＬＲＵテーブル23からそのアドレスのグループＩＤがひかれる。このグループＩＤは比較回路19においてアクセスのあったグループのアンド・ゲート11ないし14の出力と比較される。その結果不一致のときはアドレスが更新され、再びＬＲＵテーブル23から次のアドレスのグループＩＤがひかれて同様に比較回路19においてアンド・ゲート11ないし14の出力と比較され、比較回路19から一致出力が得られるまで繰り返えされる。一致出力が得られるとアンド・ゲート22からグループＩＤの一致したアドレス信号が出力される。このようにＬＲＵテーブル23でグループＩＤ＝ｉのエントリを下から探し「ＬＲＵエントリ」とする。

④ オア・ゲート10の出力の反転入力によりオア・ゲート21の出力がアンド・ゲート22に供給され、アンド・ゲート22からアドレス信号ｑが出力される。このようにＬＲＵテーブル23の最下位のエントリを「ＬＲＵエントリ」とする。

⑤ アンド・ゲート22の出力信号によりＬＲＵエントリの領域が解放され、ＬＲＵテーブルの置き換えが行われる。

第６図は本発明に使用されるＬＲＵテーブルを示す図で、各エントリにはドライブ／トラック・アドレスとグループＩＤが記入されている。一個のエントリはバッファ記憶装置３の一個の管理記憶単位に対応する。第７図は本発明に使用されるグループＩＤと割当上限値との対応表を示す図である。第８図は本発明に使用されるエクステント・テーブルを示す図で、各エントリには機番フィールドとトラック・アドレス・フィールドとグループＩＤフィールドとがある。機番フィールドとトラック・アドレス・フィールドにはバッファ記憶装置３に格納すべき記憶単位が記入されている。第９図はバッファ・テーブルを示す図で、各エントリにはバッファ記憶装置３に格納されている記憶単位のトラック・アドレスとバッファ記憶位置が記入されている。例えば、データ・セットＡ、Ｂ、Ｃ、Ｄがディスク記憶装置５にある場合は、これらのデータ・セットは、例えばアクセスの仕方、即ちアクセスがランダムに行われるのか、シーケンシャルに行われるのか等により、或いはアクセス頻度、即ち単位時間当りのアクセス回数により、或いはライト／リード比により、或いはレスポンス要求時間により、データ・セットＡとＢは同グループに、データ・セットＣとＤは別グループにというようにグループ分けが行われる。

このように各グループの占有領域量が上限値に達していない時の動作は単純ＬＲＵと同じであるが、あるグループのアクセスが膨んで新規のステージングが大量に発生すると、そのグループの占有領域量は上限値に到達する。その時には、そのグループの占有量がこれ以上増えないようにそのグループ内でＬＲＵが行われる。その後、他グループのステージングが増加してくると、このグル

ープのトラックが追い出され再び単純ＬＲＵの動作に戻る。

本発明によれば、トラックをいくつかのグループに分けるが、ＬＲＵテーブル自体は一つである。但し各グループに対して割当ての上限値を設ける。上限値はトラック数でも良いし、領域の管理単位（ブロック）でも良い。各グループの上限値の合計は実際のバッファ容量を越えていても良い。上限値はグループ内トラックによる過占有状態の抑止のためにあり、ＬＲＵテーブルをグループ毎に持つこととは異なる。

各グループの上限値の合計をバッファ容量に等しいようにしておくと、グループ毎のＬＲＵと同じことになる。その場合、領域の無駄は生ずるかもしれないが、グループとしていつも一定の領域量を確保することができる。

(6) 発明の効果

以上の説明から明らかなように、本発明によれば、定常状態では単純なＬＲＵであるが、あるグループによる占有量の増加に対して歯止めをかけることができる。従ってアクセスの性質によってグループ分けして、グループ毎に適切な上限値を設定して占有領域量を制限でき、他グループの必要なトラックが追い出されることはなく、領域の無駄が生じないのでディスク・キャッシュとして安定した性能が期待できる。

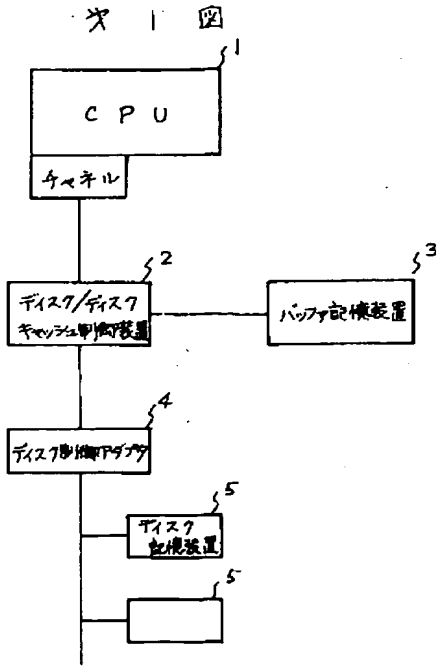
4. 図面の簡単な説明

第１図はディスク・キャッシュ・サブシステムを説明するブロック図、第２図はディスク・キャッシュ・サブシステムの処理を説明するフロー・チャート、第３図は従来のＬＲＵテーブルを示す図、第４図は本発明に使用されるハードウェアの一実施例を示すブロック図、第５図は本発明の一実施例を説明するフロー・チャート、第６図は本発明に使用されるＬＲＵテーブルを示す図、第７図は本発明に使用されるグループＩＤと割当上限値との対応表を示す図、第８図は本発明に使用されるエクステント・テーブルを示す図、第９図はバッファ・テーブルを示す図である。

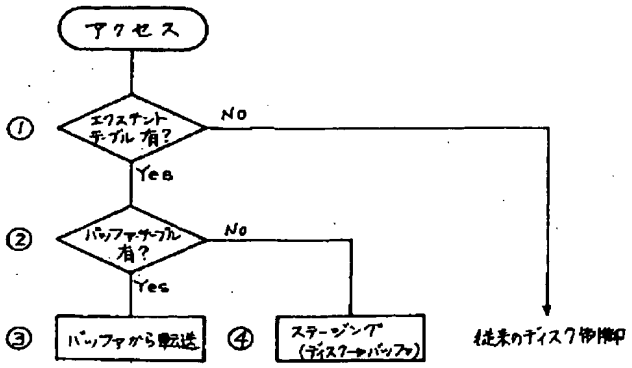
１…ホスト・コンピュータ、２…ディスク／ディスク・キャッシュ制御装置、３…バッファ記憶装置、４…ディスク制御アダプタ、５…ディスク記憶装置、６，11ないし18と22…アンド・ゲート、７…レジスタ、８…カウンタ、９と19…比較回路、10と21…オア・ゲート、20…アドレス・レジスタ、23…ＬＲＵテーブル。

特許出願人　富士通株式会社
代理人弁理士　京谷四郎

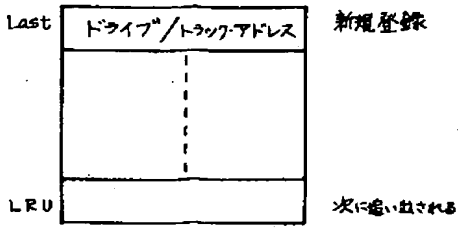
第１図



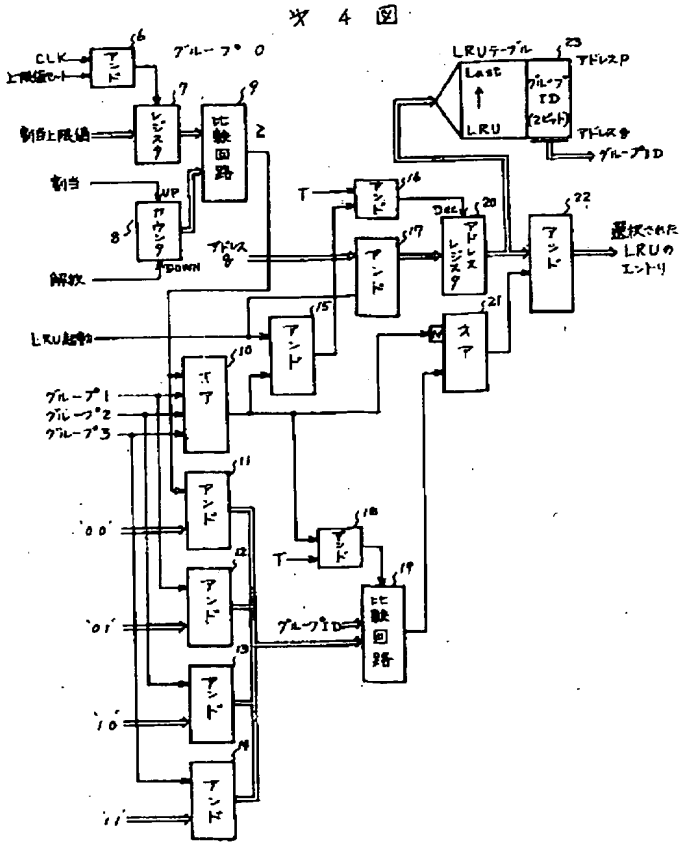
第2図


アクセス
①
エクステントテーブル有?
No
Yes
②
バッファテーブル有?
No
Yes
③
バッファから転送
④
ステージング
(ディスクキャッシュ)
従来のディスク制御部

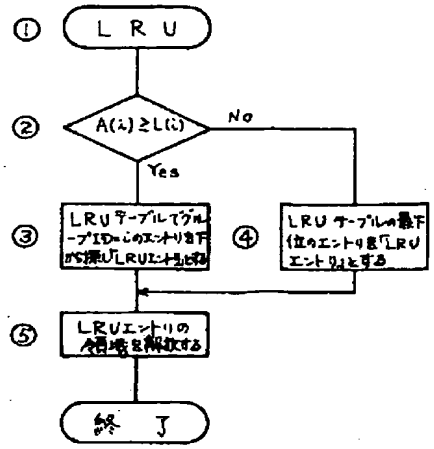
第3図


Last
ドライブ/トラック・アドレス
新規登録
LRU
次に追い出される

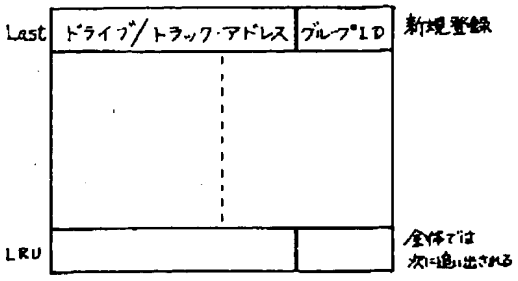
第4図


CLK
上限値セット
制御上限値
割当
解放
UP
DOWN
LRU起動
グループ0
グループ1
グループ2
グループ3
'00'
'01'
'10'
'11'
T
LRUテーブル
Last
LRU
グループID (2ビット)
アドレスP
アドレスQ
グループID
アドレス
Dec
選択されたLRUのエントリ
アンド
オア
レジスタ
比較回路
アドレスレジスタ

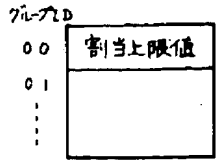
第 5 図



第 6 図



第 7 図



第 8 図

機番	トラック・アドレス from to	グループID

第 9 図

トラック・アドレス	バッファ記憶位置